SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Middeldorp, Jaap Michiel.

(ii) TITLE OF INVENTION: Peptides and nucleic acid sequences related to the Epstein-Barr virus.

(iii) NUMBER OF SEQUENCES: 22

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Organon Teknika Corporation
Biotechnology Research Institute
(B) STREET: 1330-A Piccard Drive
(C) CITY: Rockville
(D) STATE: Maryland
(E) COUNTRY: USA
(F) ZIP: 20850-4377

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: Patentin Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: EP 92200721.6
(B) FILING DATE: 13-MAR-1992
(C) CLASSIFICATION

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Bobrowicz, Donna
(B) REGISTRATION NUMBER: 32,196

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 538 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CATGATGGCA CGCCGGCTGC CCAAGCCCAC CCTCCAGGGG AGGCTGGAGG CGGATTTTCC 60

AGACAGTCCC CTGCTTCCTA AATTTCAAGA GCTGAACCAG AATAATCTCC CCAATGATGT 120

TTTTCGGGAG GCTCAAAGAA GTTACCTGGT ATTTCTGACA TCCCAGTTCT GCTACGAAGA 180

GTACGTGCAG AGGACTTTTG GGGTGCCTCG GCGCCAACGC GCCATAGACA AGAGGCAGAG 240

AGCCAGTGTG GCTGGGGCTG GTGCTCATGC ACACCTTGGC GGGTCATCCG CCACCCCCGT 300

CCAGCAGGCT CAGGCCGCCG CATCCGCTGG GACCGGGGCC TTGGCATCAT CAGCGCCGTC 360

CACGGCCGTA GCCCAGTCCG CGACCCCCTC TGTTTCTTCA TCTATTAGCA GCCTCCGGGC 420

CGCGACTTCG GGGGCGACTG CCGCCGCCTC CGCCGCCGCA GCCGTCGATA CCGGGTCAGG 480

TGGCGGGGGA CAACCCCACG ACACCGCCCC ACGCGGGGCA CGTAAGAAAC AGTAGCCC 538

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 176 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Ala Arg Arg Leu Pro Lys Pro Thr Leu Gln Gly Arg Leu Glu Ala
1 5 10 15

Asp Phe Pro Asp Ser Pro Leu Leu Pro Lys Phe Gln Glu Leu Asn Gln
20 25 30

Asn Asn Leu Pro Asn Asp Val Phe Arg Glu Ala Gln Arg Ser Tyr Leu
35 40 45

Val Phe Leu Thr Ser Gln Phe Cys Tyr Glu Glu Tyr Val Gln Arg Thr
50 55 60

Phe Gly Val Pro Arg Arg Gln Arg Ala Ile Asp Lys Arg Gln Arg Ala
65 70 75 80

Ser Val Ala Gly Ala Gly Ala His Ala His Leu Gly Gly Ser Ser Ala
85 90 95

Thr Pro Val Gln Gln Ala Gln Ala Ala Ala Ser Ala Gly Thr Gly Ala
100 105 110

Leu Ala Ser Ser Ala Pro Ser Thr Ala Val Ala Gln Ser Ala Thr Pro
115 120 125

Ser Val Ser Ser Ser Ile Ser Ser Leu Arg Ala Ala Thr Ser Gly Ala
130 135 140

Thr Ala Ala Ala Ser Ala Ala Ala Ala Val Asp Thr Gly Ser Gly Gly
145 150 155 160

Gly Gly Gln Pro His Asp Thr Ala Pro Arg Gly Ala Arg Lys Lys Gln
165 170 175

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1038 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATGCTATCAG GTAACGCAGG AGAAGGAGCA ACAGCCTGCG GAGGTTCGGC CGCCGCGGGC 60

CAGGACCTCA TCAGCGTCCC CCGCAACACC TTTATGACAC TGCTTCAGAC CAACCTGGAC 120

AACAAACCGC CGAGGCAGAC CCCGCTACCC TACGCGGCCC CGCTGCCCCC CTTTTCCCAC 180

CAGGCAATAG CCACCGCGCC TTCCTACGGT CCTGGGGCCG GAGCGGTCGC CCCGGCCGGC 240

GGCTACTTTA CCTCCCCAGG AGGTTACTAC GCCGGGCCCG CGGGCGGGGA CCCGGGTGCC 300

TTCTTGGCGA TGGACGCTCA CACCTACCAC CCCCACCCAC ACCCCCCTCC GGCCTACTTT 360

GGCTTGCCGG GCCTCTTTGG CCCCCCTCCA CCCGTGCCTC CTTACTACGG ATCCCACTTG 420

CGGGCAGACT ACGTCCCCGC TCCCTCGCGA TCCAACAAGC GGAAAAGAGA CCCCGAGGAG 480

GATGAAGAAG GCGGGGGGCT ATTCCCGGGG GAGGACGCCA CCCTCTACCG CAAGGACATA 540

GCGGGCCTCT CCAAGAGTGT GAATGAGTTA CAGCACACGC TACAGGCCCT GCGCCGGGAG 600

ACGCTGTCCT ACGGCCACAC CGGAGTCGGA TACTGCCCCC AGCAGGGCCC CTGCTACACC 660

CACTCGGGGC CTTACGGATT TCAGCCTCAT CAAAGCTACG AAGTGCCCAG ATACGTCCCT 720

CATCCGCCCC CACCACCAAC TTCTCACCAG GCAGCTCAGG CGCAGCCTCC ACCCCCGGGC 780

ACACAGGCCC CCGAAGCCCA CTGTGTGGCC GAGTCCACGA TCCCTGAGGC GGGAGCAGCC 840

GGGAACTCTG GACCCCGGGA GGACACCAAC CCTCAGCAGC CCACCACCGA GGGCCACCAC 900

CGCGGAAAGA AACTGGTGCA GGCCTCTGCG TCCGGAGTGG CTCAGTCTAA GGAGCCCACC 960

ACCCCCAAGG CCAAGTCTGT GTCAGCCCAC CTCAAGTCCA TCTTTTGCGA GGAATTGCTG 1020

AATAAACGCG TGGCTTGA 1038

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 345 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met Leu Ser Gly Asn Ala Gly Glu Gly Ala Thr Ala Cys Gly Gly Ser
1 5 10 15

Ala Ala Ala Gly Gln Asp Leu Ile Ser Val Pro Arg Asn Thr Phe Met
20 25 30

Thr Leu Leu Gln Thr Asn Leu Asp Asn Lys Pro Pro Arg Gln Thr Pro
35 40 45

Leu Pro Tyr Ala Ala Pro Leu Pro Pro Phe Ser His Gln Ala Ile Ala
50 55 60

Thr Ala Pro Ser Tyr Gly Pro Gly Ala Gly Ala Val Ala Pro Ala Gly
65 70 75 80

Gly Tyr Phe Thr Ser Pro Gly Gly Tyr Tyr Ala Gly Pro Ala Gly Gly
85 90 95

Asp Pro Gly Ala Phe Leu Ala Met Asp Ala His Thr Tyr His Pro His
100 105 110

Pro His Pro Pro Pro Ala Tyr Phe Gly Leu Pro Gly Leu Phe Gly Pro
115 120 125

Pro Pro Pro Val Pro Pro Tyr Tyr Gly Ser His Leu Arg Ala Asp Tyr
130 135 140

Val Pro Ala Pro Ser Arg Ser Asn Lys Arg Lys Arg Asp Pro Glu Glu
145 150 155 160

Asp Glu Glu Gly Gly Gly Leu Phe Pro Gly Glu Asp Ala Thr Leu Tyr

```
                    165                 170                 175

Arg Lys Asp Ile Ala Gly Leu Ser Lys Ser Val Asn Glu Leu Gln His
                180                 185                 190

Thr Leu Gln Ala Leu Arg Arg Glu Thr Leu Ser Tyr Gly His Thr Gly
        195                 200                 205

Val Gly Tyr Cys Pro Gln Gln Gly Pro Cys Tyr Thr His Ser Gly Pro
    210                 215                 220

Tyr Gly Phe Gln Pro His Gln Ser Tyr Glu Val Pro Arg Tyr Val Pro
225                 230                 235                 240

His Pro Pro Pro Pro Pro Thr Ser His Gln Ala Ala Gln Ala Gln Pro
                245                 250                 255

Pro Pro Pro Gly Thr Gln Ala Pro Glu Ala His Cys Val Ala Glu Ser
            260                 265                 270

Thr Ile Pro Glu Ala Gly Ala Ala Gly Asn Ser Gly Pro Arg Glu Asp
        275                 280                 285

Thr Asn Pro Gln Gln Pro Thr Thr Glu Gly His His Arg Gly Lys Lys
    290                 295                 300

Leu Val Gln Ala Ser Ala Ser Gly Val Ala Gln Ser Lys Glu Pro Thr
305                 310                 315                 320

Thr Pro Lys Ala Lys Ser Val Ser Ala His Leu Lys Ser Ile Phe Cys
                325                 330                 335

Glu Glu Leu Leu Asn Lys Arg Val Ala
            340                 345
```

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```
Ala Val Asp Thr Gly Ser Gly Gly Gly Gly Gln Pro His Asp Thr Ala
                5                   10                  15

Pro Arg Gly Ala Arg Lys Lys Gln
            20
```

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Ser Thr Ala Val Ala Gln Ser Ala Thr Pro Ser Val Ser Ser Ser Ile
5 10 15

Ser Ser Leu Arg Ala Ala Thr Ser Gly Ala Thr Ala Ala Ala
20 25

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Gly Val Pro Arg Arg Gln Arg Ala Ile Asp Lys Arg Gln Arg Ala
5 10 15

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Gly Gln Pro His Asp Thr Ala Pro Arg Gly Ala Arg Lys Lys Gln
5 10 15

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Thr Ala Val Ala Gln Ser Ala Thr Pro Ser Val Ser
5 10

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Pro Ser Val Ser Ser Ser Ile Ser Ser Leu Arg Ala
5 10

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

Ser Val Ser Ser Ser Ile Ser Ser Leu Arg Ala Ala
5 10

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Ser Ser Ser Ile Ser Ser Leu Arg Ala Ala Thr Ser
5 10

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

Ser Ile Ser Ser Leu Arg Ala Ala Thr Ser Gly Ala
5 10

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

Ile Ser Ser Leu Arg Ala Ala Thr Ser Gly Ala Thr
5 10

(2) INFORMATION FOR SEQ ID NO: 15

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15

Arg Ala Ala Thr Ser Gly Ala Thr Ala Ala Ala Ser
5 10

(2) INFORMATION FOR SEQ ID NO: 16

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Ala Ala Val Asp Thr Gly Ser Gly Gly Gly Gly Gln
5 10

(2) INFORMATION FOR SEQ ID NO: 17

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

Ala Val Asp Thr Gly Ser Gly Gly Gly Gly Gln Pro
5 10

(2) INFORMATION FOR SEQ ID NO: 18

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Val Asp Thr Gly Ser Gly Gly Gly Gly Gln Pro His
5 10

(2) INFORMATION FOR SEQ ID NO: 19

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

Asp Thr Gly Ser Gly Gly Gly Gly Gln Pro His Asp
5 10

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

Gly Gly Gly Gln Pro His Asp Thr Ala Pro Arg Gly
5 10

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

Gly Gln Pro His Asp Thr Ala Pro Arg Gly Ala Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Epstein-Barr virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

Gln Pro His Asp Thr Ala Pro Arg Gly Ala Arg Lys
5 10